Edge and Region Analysis for Digital Image Data

ROBERT M. HARALICK

Department of Electrical Engineering and Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

In this paper we provide a unified view of edge and region analysis. Our framework is based on the sloped-facet model which assumes that regions of image segments are maximal areas which are sloped planes. Edge strength between two adjacent pixels is measured by the F statistic used to test the significance of the difference between the parameters of the best-fitting sloped neighborhoods containing each of the given pixels. Edges are declared to exist at locations of local maxima in the F-statistic edge strength picture. We show that this statistically optimum procedure in essence scales the edge strength statistic of many popular edge operators by an estimate of the image noise. Such a scaling makes optimum detection possible by a fixed threshold procedure.

1. INTRODUCTION

Edge detection and region growing are two areas of image analysis which are opposite in emphasis but identical at heart. Edges obviously occur at bordering locations of two adjacent regions which are significantly different. Regions are maximal areas having similar attributes. If we could do region analysis, then edges could be declared at the borders of all regions. If we could do edge detection, regions would be the areas surrounded by the edges. Unfortunately, we tend to have trouble doing either: edge detectors are undoubtedly noisy and region growers often grow too far.

In this paper we give an even-handed treatment of both. Edges will not occur at locations of high differences. Rather, they will occur at locations of high differences between the parameters of sufficiently homogeneous areas. Regions will not be declared as just areas of similar value of gray tone. They will occur at connected areas where resolution cells yield minimal differences of region parameters, where minimal means smallest among a set of resolution cell groupings. In essence we will see that edge detection and region analysis are identical problems that can be solved with the same procedure.

Because the framework we wish to present tends to unify some of the popular techniques, the paper is organized to first give a description of our framework and then to describe related techniques discussed in the literature in terms of our framework.

ISBN 0-12-597320-9

2. THE SLOPED-FACET MODEL

The digital image g is a function from the Cartesian product of row and column index sets into the reals. The sloped-facet assumption is a restriction on the nature of the function g for the ideal image (one having no defocusing or noise). The restriction is that the domain of g can be partitioned into connected sets $\Pi = \{\Pi_1, \ldots, \Pi_M\}$ such that for each connected set $\Pi_m \in \Pi$.

(1) $(r, c) \in \Pi_m$ implies that for some K-pixel neighborhood N containing (r, c), $N \subseteq \Pi_m$,
(2) $(r, c) \in \Pi_m$ implies $g(r, c) = \alpha_m r + \beta_m c + \gamma_m$.

Condition (1) requires that the partition Π consists of connected sets each of large enough and smooth enough shape. For example, if $K = 9$ and the only neighborhoods we consider are 3×3, than each set Π_m must be no thinner in any place than 3×3. If Π_m has holes, they must be surrounded everywhere by pixels in 3×3 neighborhoods which are entirely contained in Π_m.

Condition (2) requires that the gray tone surface defined on Π_m be a sloped plane. This constraint could obviously be generalized to include higher-order polynomials. It is, of course, more general than the piecewise constant surface implicitly assumed by other techniques.

The fact that the parameters α and β determine the value of the slope in any direction is well known. For a planar surface of the form

$$g(r, c) = \alpha r + \beta c + \gamma$$

the value of the slope at an angle θ to the row axis is given by the directional derivative of g in the direction θ. Since α is the partial derivative of g with respect to r and β is the partial derivative of g with respect to c, the value of the slope at angle θ is $\alpha \cos \theta + \beta \sin \theta$. Hence, the value of the slope at any direction is an appropriate linear combination of the values for α and β. The angle θ which maximizes this value satisfies

$$\cos \theta = \frac{\alpha}{(\alpha^2 + \beta^2)^{\frac{1}{2}}} \quad \text{and} \quad \sin \theta = \frac{\beta}{(\alpha^2 + \beta^2)^{\frac{1}{2}}}$$

and the gradient which is the value of the slope in the steepest direction is $(\alpha^2 + \beta^2)^{\frac{1}{2}}$.

The sloped-facet assumption can avoid some of the problems inherent in edge detectors or region growers. Consider, for example, two piecewise linear surfaces meeting at a V-junction. A typical step edge detector applied at the V-junction would tend to find the average gray tone to the left of the junction equal to the average gray tone to the right of the junction. A region grower, for the same reason, if approaching the junction from the left is likely to grow somewhat into the part to the right of the junction before it realizes that there may be significant gray tone difference. Consider also a simple sloped surface. A typical step edge detector applied any place along this surface might declare an edge because the gray tone average to the detector's left is surely different than the average to its right.

The sloped-facet model is an appropriate one for either the flat-world or sloped-world assumption. In the flat world each ideal region is constant in gray tone. Hence, all edges are step edges. The observed image taken in an ideal flat world is a defocused version of the ideal piecewise constant image with the addition of some random noise. The defocusing changes all step edges to sloped edges. The edge detection problem is one of determining whether the observed noisy slope has a gradient significantly higher than one which could have been caused by the noise alone. Edge boundaries are declared in the middle of all significantly sloped regions.

In the sloped facet world, each ideal region has a gray tone surface which is a sloped plane. Edges are places of either discontinuity in gray tone or derivative of gray tone. The observed image is the ideal image with noise added and no defocusing. To determine if there is an edge between two pixels, we first determine the best slope fitting neighborhood for each of the pixels. Edges are declared at locations having significantly different planes on either side of them. In the sloped facet model, edges surrounding regions having significantly sloped surfaces may be the boundaries of an edge region. The determination of whether a sloped region is an edge region or not may depend on the significance and magnitude of the slope as well as the semantics of the image.

In either the noisy defocused flat world or the noisy sloped world we are faced with the problem of estimating the parameters of a sloped surface for a given neighborhood and then calculating the significance of the difference of the estimated slope from a zero slope or calculating the significance of the difference of the estimated slopes of two adjacent neighborhoods. To do this we proceed in a classical manner. We will use a least-squares procedure to estimate parameters and we will measure the strength of any difference by an appropriate F statistic.

3. SLOPED FACET PARAMETER ESTIMATION AND SIGNIFICANCE MEASURE

We employ a least-squares procedure to estimate the parameters of the slope model for a rectangular region whose row index set is R and whose column index set is C. We assume that for each $(r, c) \in R \times C$,

$$g(r, c) = \alpha r + \beta c + \gamma + \eta(r, c)$$

where η is a random variable indexed on $R \times C$ which represents noise. We will assume that η is noise having mean 0 and variance σ^2 and that the noise for any two pixels is independent.

The least-squares procedure determines an $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ which minimize

$$\epsilon^2 = \sum_{r \in R} \sum_{c \in C} [\hat{\alpha} r + \hat{\beta} c + \hat{\gamma} - g(r, c)]^2.$$

Taking the partial derivatives of ϵ^2 and setting them to zero results in

$$\begin{bmatrix} \dfrac{\partial \epsilon^2}{\partial \hat{\alpha}} \\ \dfrac{\partial \epsilon^2}{\partial \hat{\beta}} \\ \dfrac{\partial \epsilon^2}{\partial \hat{\gamma}} \end{bmatrix} = 2 \sum_{r \in R} \sum_{c \in C} (\hat{\alpha} r + \hat{\beta} c + \hat{\gamma} - g(r, c)) \begin{bmatrix} r \\ c \\ 1 \end{bmatrix} = 0. \tag{1}$$

Without loss of generality, we choose our coordinate system $R \times C$ so that the center of the neighborhood $R \times C$ has coordinates (0, 0). When the number of rows and columns is odd, the center pixel, therefore, has coordinates (0, 0). When the number of rows and columns is even, there can be no one center pixel and the point where the four center pixels meet has coordinates (0, 0).

The symmetry in the chosen coordinate system leads to

$$\sum_{r \in R} r = \sum_{c \in C} c = 0.$$

Hence, Eq. (1) reduces to the system of three decoupled equations

$$\sum_{r \in R} \sum_{c \in C} \hat{\alpha} r^2 = \sum_{r \in R} \sum_{c \in C} r g(r, c),$$

$$\sum_{r \in R} \sum_{c \in C} \hat{\beta} c^2 = \sum_{r \in R} \sum_{c \in C} c g(r, c),$$

$$\sum_{r \in R} \sum_{c \in C} \hat{\gamma} = \sum_{r \in R} \sum_{c \in C} g(r, c).$$

Solving for $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ we obtain

$$\begin{aligned} \hat{\alpha} &= \sum_{r \in R} \sum_{c \in C} r g(r, c) \Big/ \sum_{r \in R} \sum_{c \in C} r^2, \\ \hat{\beta} &= \sum_{r \in R} \sum_{c \in C} c g(r, c) \Big/ \sum_{r \in R} \sum_{c \in C} c^2, \\ \hat{\gamma} &= \sum_{r \in R} \sum_{c \in C} g(r, c) \Big/ \sum_{r \in R} \sum_{c \in C} 1. \end{aligned} \tag{2}$$

Replacing $g(r, c)$ by $\alpha r + \beta c + \gamma + \eta(r, c)$ and reducing the equations will allow us to explicitly see the dependence of $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ on the noise. We obtain

$$\hat{\alpha} = \alpha + \Big(\sum_r \sum_c r\eta(r, c) \Big/ \sum_r \sum_c r^2\Big),$$

$$\hat{\beta} = \beta + \Big(\sum_r \sum_c c\eta(r, c) \Big/ \sum_r \sum_c c^2\Big),$$

$$\hat{\gamma} = \gamma + \Big(\sum_r \sum_c \eta(r, c) \Big/ \sum_r \sum_c 1\Big).$$

From this it is apparent that $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are unbiased estimators for α, β, and γ, respectively, and have variances

$$V[\hat{\alpha}] = \sigma^2 / \sum_{r\in R} \sum_{c\in C} r^2,$$

$$V[\hat{\beta}] = \sigma^2 / \sum_{r\in R} \sum_{c\in C} c^2,$$

$$V[\hat{\gamma}] = \sigma^2 / \sum_{r\in R} \sum_{c\in C} 1.$$

Normally distributed noise implies that $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are also normally distributed. The independence of the noise implies that $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are independent since they are normal and that

$$E[(\hat{\alpha} - \alpha)(\hat{\beta} - \beta)] = E[(\hat{\alpha} - \alpha)(\hat{\gamma} - \gamma)] = E[(\hat{\beta} - \beta)(\hat{\gamma} - \gamma)] = 0$$

as a straightforward calculation shows.
Examining the total squared error ϵ^2 we find that

$$\begin{aligned}
\epsilon^2 &= \sum_{r\in R} \sum_{c\in C} [(\hat{\alpha}r + \hat{\beta}c + \hat{\gamma}) - (\alpha r + \beta c + \gamma + \eta(r, c))]^2 \\
&= \sum_{r\in R} \sum_{c\in C} [(\hat{\alpha} - \alpha)^2 r^2 + (\hat{\beta} - \beta)^2 c^2 + (\hat{\gamma} - \gamma)^2 + \eta^2(r, c) \\
&\qquad - 2(\hat{\alpha} - \alpha) r\eta(r, c) - 2(\hat{\beta} - \beta) c\eta(r, c) - 2(\hat{\gamma} - \gamma)\eta(r, c)].
\end{aligned}$$

Using the fact that

$$(\hat{\alpha} - \alpha) = \sum_{r\in R} \sum_{c\in C} r\eta(r, c) / \sum_r \sum_c r^2,$$

$$(\hat{\beta} - \beta) = \sum_r \sum_c c\eta(r, c) / \sum_r \sum_c c^2,$$

$$(\hat{\gamma} - \gamma) = \sum_{r\in R} \sum_c \eta(r, c) / \sum_r \sum_c 1$$

we may substitute into the last three terms for ϵ^2 and obtain after simplification

$$\epsilon^2 = \sum_r \sum_c \eta^2(r, c) - (\hat{\alpha} - \alpha)^2 \sum_r \sum_c r^2 - (\hat{\beta} - \beta)^2 \sum_r \sum_c c^2 - (\hat{\gamma} - \gamma)^2 \sum_r \sum_c 1.$$

Now notice that

$$\sum_r \sum_c \eta^2(r, c)$$

is the sum of the squares of

$$\sum_r \sum_c 1$$

independently distributed normal random variables. Hence,

$$\sum_r \sum_c \eta^2(r, c) / \sigma^2$$

is distributed as a chi-squared variate with

$$\sum_r \sum_c 1$$

degrees of freedom. Because, $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are independent normals,

$$((\hat{\alpha} - \alpha)^2 \sum_r \sum_c r^2 + (\hat{\beta} - \beta)^2 \sum_r \sum_c c^2 + (\hat{\gamma} - \gamma)^2 \sum_r \sum_c 1)/\sigma^2$$

is distributed as a chi-squared variate with 3 degrees of freedom. Therefore, ϵ^2/σ^2 is distributed as a chi-squared variate with

$$\sum_r \sum_c 1 - 3$$

degrees of freedom.

From this it follows that to test the hypothesis of no edge for the flat-world assumption, $\alpha = \beta = 0$, we use the ratio

$$F = ((\hat{\alpha}^2 \sum_r \sum_c r^2 + \hat{\beta}^2 \sum_r \sum_c c^2)/2)/(\epsilon^2/(\sum_r \sum_c 1 - 3)),$$

which has an F distribution with

$$2, \sum_r \sum_c 1 - 3$$

degrees of freedom and reject the hypothesis for large values of F.

Notice that F may be regarded as a significance or reliability measure associated with the existence of a nonzero sloped region in the domain $R \times C$. It is essentially proportional to the squared gradient of the region normalized by

$$\epsilon^2/(\sum_r \sum_c 1 - 3)$$

which is a random variable whose expected value is σ^2, the variance of the noise.

EXAMPLE 1. Consider the following 3×3 region:

3	5	9	
4	7	7	observed.
0	3	7	

Then $\hat{\alpha} = -1.17$, $\hat{\beta} = 2.67$, and $\hat{\gamma} = 5.00$. The estimated gray tone surface is given by $\hat{\alpha}r + \hat{\beta}c + \hat{\gamma}$ and is

3.50	6.17	8.83	
2.33	5.00	7.67	estimated.
1.17	3.83	6.5	

The difference between the estimated and the observed surfaces is the error and it is

0.50	1.17	−0.17	
−1.67	−2.00	0.67	error.
1.17	0.83	−0.50	

From this we can compute the squared error $\epsilon^2 = 11.19$. The F statistic is then

$$\frac{[(-1.17)^2 \cdot 6 + (2.67)^2 \cdot 6]/2}{11.19/6} = 13.67.$$

If we were compelled to make a hard decision about the significance of the observed slope in the given 3×3 region, we would probably call it a nonzero sloped region since the probability of a region with true zero slope giving an $F_{2,6}$ statistic of value less than 10.6 is 0.99. 13.67 is greater than 10.6 so we are assured that the probability of calling the region a nonzero sloped region when it is in fact a zero sloped region is much less than 1%. The statistically oriented reader will recognize the test as a 1% significance level test.

EXAMPLE 2. We proceed just as in Example 1 for the following 3×3 region:

1	3	11	
6	11	7	observed.
−4	1	9	

Then $\hat{\alpha} = 1.5$, $\hat{\beta} = 4.0$, and $\hat{\gamma} = 5.0$. The estimated gray tone surface is

2.5	6.5	10.5	
1	5	9	estimated
−0.5	3.5	7.5	

and the error surface is

1.5	3.5	−0.5	
−5	−6	2	error.
3.5	2.5	−1.5	

From this we compute an F statistic of 3.27 and hence we call it a zero sloped region at the 1% significance level.

For the sloped-facet world our problem is not whether the true slope of a region is zero; rather, it is determining whether two regions are part of the same sloped surface. To do this we are naturally led to examine the differences between the parameters for the estimated sloped-plane surfaces.

For simplicity, we assume that the two regions 1 and 2 are identically sized mutually exclusive rectangular regions. Let $\hat{\alpha}_1$, $\hat{\beta}_1$, and $\hat{\gamma}_1$ be the estimated parameters for region 1. Let $(\Delta r, \Delta c)$ be the coordinates of the center of region 2 relative to the center of region 1. Let $\hat{\alpha}_2$, $\hat{\beta}_2$, and $\hat{\gamma}_2$ be the estimated parameters for region 2. Then under the hypothesis that

$$\alpha_1 = \alpha_2 \quad \text{and} \quad \beta_1 = \beta_2,$$

$$(1/2^{\frac{1}{2}})(\hat{\alpha}_1 - \hat{\alpha}_2)(\sum_r \sum_c r^2)^{\frac{1}{2}} \quad \text{and} \quad (1/2^{\frac{1}{2}})(\hat{\beta}_1 - \hat{\beta}_2)(\sum_r \sum_c c^2)^{\frac{1}{2}}$$

each have a normal distribution with mean 0 and variance σ^2.

Due to the fact that the gray tone surfaces are sloped the hypothesis $\gamma_1 = \gamma_2$ is inappropriate. Instead, we must adjust the average height for each surface to account for the sloped rise or fall as we travel from the center of region 1 to the center of region 2 and test the equality of the adjusted heights. To do this we choose a place halfway between the centers of the two regions. Since each region's center had relative coordinates (0, 0), the coordinates of the halfway location from region 1 is $(\Delta r/2, \Delta c/2)$ and the coordinates of the halfway location from region 2 is $(-\Delta r/2, -\Delta c/2)$. The true height of the gray tone surface at these locations is

$$\alpha_1 \Delta r/2 + \beta_1 \Delta c/2 + \gamma_1 \qquad \text{and} \qquad \alpha_2(-\Delta r/2) + \beta_2(-\Delta c/2) + \gamma_2$$

and the hypothesis that the regions are part of the same sloped surface would imply

$$(\alpha_1 + \alpha_2)\frac{\Delta r}{2} + (\beta_1 + \beta_2)\frac{\Delta c}{2} + \gamma_1 - \gamma_2 = 0.$$

Under this hypothesis the statistic

$$(\hat{\alpha}_1 + \hat{\alpha}_2)\frac{\Delta r}{2} + (\hat{\beta}_1 + \hat{\beta}_2)\frac{\Delta c}{2} + (\hat{\gamma}_1 - \hat{\gamma}_2)$$

has a normal distribution with mean 0 and variance

$$\sigma^2[2((\Delta r/2)^2/\sum_r \sum_c r^2) + 2((\Delta c/2)^2/\sum_r \sum_c c^2) + (2/\sum_r \sum_c 1)].$$

After noting that under the same sloped-surface assumption

$$E[(\hat{\alpha}_1 + \hat{\alpha}_2)(\hat{\alpha}_1 - \hat{\alpha}_2)] = E[(\hat{\beta}_1 + \hat{\beta}_2)(\hat{\beta}_1 - \hat{\beta}_2)] = 0$$

we find that an appropriate statistic to measure the significance of the departure from the same-surface hypothesis is

$$\left\{\tfrac{1}{2}(\hat{\alpha}_1 - \hat{\alpha}_2)^2 \sum_r \sum_c r^2 + \tfrac{1}{2}(\hat{\beta}_1 - \hat{\beta}_2)^2 \sum_r \sum_c c^2 + \left(\frac{[(\hat{\alpha}_1 + \hat{\alpha}_2)(\Delta r/2) + (\hat{\beta}_1 + \hat{\beta}_2)(\Delta c/2) + \hat{\gamma}_1 - \hat{\gamma}_2]^2}{2[(\Delta r/2)^2/\sum_r \sum_c r^2 + (\Delta c/2)^2/\sum_r \sum_c c^2 + 1/\sum_r \sum_c 1]}\right)\right\} \Big/ 3 \Big/ [(\epsilon_1^2 + \epsilon_2^2)/(2\sum_r \sum_c 1 - 6)]$$

EXAMPLE 3. We proceed just as in Examples 1 and 2 for the following pair of 3×3 regions:

17	15	12				
20	16	13	18	14	12	
21	18	15	20	15	11	observed.
			19	19	16	

Here, $\hat{\alpha}_1 = \hat{\alpha}_2 = 1.67$, $\hat{\beta}_1 = \hat{\beta}_2 = -3$, $\hat{\gamma}_1 = 16.33$, and $\hat{\gamma}_2 = 16.00$. The estimated gray tone surface is

17.33	14.33	11.33				
19	16	13	17.55	14.55	11.55	
20.67	17.67	14.67	19.22	16.22	13.22	estimated
			20.88	17.88	14.88	

and the error surface is

$$\begin{array}{rrr} -0.67 & 0.33 & 0.67 \\ -1.0 & 1.0 & 2.0 \\ 1.67 & -1.33 & -1.33 \\ & & \end{array} \qquad \begin{array}{rrr} & & \\ -0.44 & -0.44 & 0.55 \\ 2.22 & 0.22 & -1.77 \\ -2.12 & 0.88 & 0.88 \end{array} \quad \text{error.}$$

From this we compute an F statistic for the hypothesis that the two regions are part of the same sloped surface:

$$F = \left(\frac{(1.67 + (-9) + 0.33)^2}{2(0.5^2/6 + 1.5^2/6 + 1/a)} \Big/ 3\right) \Big/ (14.64/12) = 12.68$$

At the 1% significance level we would reject this hypothesis. Note that an edge detector comparing simple differences of average gray tones would most likely not call this an edge.

To detect whether an edge exists between a pair of neighboring pixels, we can employ the following procedure. For each pixel, examine all neighborhoods containing the pixel and determine that neighborhood whose sloped-surface fitting error is least. If the best-fitting neighborhoods for the two pixels overlap, then declare no edge. If the best-fitting neighborhoods for the two pixels have no overlap, then compute the F statistic to measure the significance of the difference between the two sloped surfaces.

4. USING THE SLOPED-FACET MODEL

To find regions we must look for connected sets of resolution cells which are surely on the same sloped surface. To find edges we must look for pairs of adjacent regions having significantly different sloped surfaces. To do edge detection and region analysis we must do both. This suggests the following way to apply the sloped-facet model. Select an appropriate sized neighborhood. Run this neighborhood over the image. For each location where the neighborhood may be placed on the image, determine the α, β, γ parameters of the sloped surface fit as well as the ϵ^2 error of the fit. Use this information to create an image in which each resolution cell has the four parameters α, β, γ, ϵ^2 from its corresponding neighborhood in the given image.

Although the sloped-surface parameters are placed in a single location, they actually apply to all the pixels in the neighborhood on the original image over which the fitting was done. Hence, on the original image, each pixel has associated with it the four parameters α, β, γ, and ϵ^2 for *each* neighborhood that contains it. From all of the neighborhoods of a pixel, one will have a parameter set having the smallest ϵ^2, that is, the best fit. Now associate with each pixel on the original image the parameters associated with its best-fitting neighborhood as well as the coordinates for this best-fitting neighborhood.

To do edge detection, for each pair of adjacent pixels compute the F statistic associated with the hypothesis that the pair of best-fitting neighborhoods corresponding to these pixels have the same sloped surface. Of course, set the F

statistic to zero if the best neighborhoods overlap. In this manner, an image of these F statistics can be created for vertical and horizontal edges. Then apply a non-maximum suppression operator (Rosenfeld and Kak [6]) in a direction orthogonal to the edge direction to determine those locations having a local maximum in the F-statistic. These locations should be the edges.

To do region growth use a clustering or grouping method. Join together each pair of neighboring pixels if their best neighborhoods overlap. If their neighborhoods do not overlap join them together if the hypothesis that the pixels' best neighborhoods are part of the same sloped surface cannot be rejected.

5. LITERATURE REVIEW

The idea of fitting linear surfaces for edge detection is not new. Roberts [1] employed an operator commonly called the Roberts gradient to determine edge strength in a 2×2 window in a blocks-world scene analysis problem. The fact that the Roberts gradient arises from a linear fit over the 2×2 neighborhood may surprise some people. To see this, let

$$\begin{matrix} a & b \\ c & d \end{matrix}$$

be the four gray tone levels in the 2×2 window whose coordinates are

$$\begin{matrix} (-\frac{1}{2}, -\frac{1}{2}) & (-\frac{1}{2}, \frac{1}{2}) \\ (\frac{1}{2}, -\frac{1}{2}) & (\frac{1}{2}, \frac{1}{2}) \end{matrix}$$

The least-squares fit for α and β is

$$\hat{\alpha} = \tfrac{1}{2}[(c + d) - (a + b)] \quad \text{and} \quad \hat{\beta} = \tfrac{1}{2}[(b + d) - (a + c)].$$

The gradient, which is the slope in the steepest direction, has magnitude $(\hat{\alpha}^2 + \hat{\beta}^2)^{\frac{1}{2}}$.

$$\begin{aligned} (\hat{\alpha}^2 + \hat{\beta}^2)^{\frac{1}{2}} &= \tfrac{1}{2}([(c + d) - (a + b)]^2 + [(b + d)^2 - (a + c)]^2)^{\frac{1}{2}} \\ &= \tfrac{1}{2}(2(a - d)^2 + 2(b - c)^2)^{\frac{1}{2}} \\ &= (2^{\frac{1}{2}}/2)((a - d)^2 + (b - c)^2)^{\frac{1}{2}} \end{aligned}$$

which is exactly $2^{\frac{1}{2}}/2$ times the Roberts gradient.

Prewitt [2] used a quadratic fitting surface to estimate the parameters α and β in a 3×3 window for automatic leukocyte cell scan analysis. The resulting values for $\hat{\alpha}$ and $\hat{\beta}$ are the same for the quadratic or linear fit. O'Gorman and Clowes [8] also used a 3×3 window and linear fit. Brooks [7] and Hummel [16] discussed the general fitting idea. Meró and Vamos [10] used the fitting idea to find lines on a binary image. Hueckel [11] used the fitting idea with low-frequency polar-form Fourier basis functions on a circular disk in order to detect step edges.

Figure 1 illustrates the windows for computing $\hat{\alpha}$ and $\hat{\beta}$ for neighborhood sizes of 2, 3, 4, and 5. Note that for the larger windows, the pixels on the edges of the window have higher weights. This means that the idea of using the difference

$$\alpha \qquad\qquad \beta \qquad\qquad \gamma$$

$$\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \qquad \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\frac{1}{6}\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad \frac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \qquad \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\frac{1}{20}\begin{bmatrix} -3 & -3 & -3 & -3 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 3 & 3 & 3 & 3 \end{bmatrix} \qquad \frac{1}{20}\begin{bmatrix} -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \end{bmatrix} \qquad \frac{1}{16}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\frac{1}{50}\begin{bmatrix} -2 & -2 & -2 & -2 & -2 \\ -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix} \qquad \frac{1}{50}\begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \qquad \frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

FIG. 1. The masks for estimating α, β, and γ for 2×2, 3×3, 4×4, and 5×5 neighborhoods.

of equally weighted averages to measure the slope in a given direction is incorrect. However, under the condition that $\alpha = \beta = 0$ an appropriate statistic to test whether two neighborhoods are on the same constant surface is an F statistic of the form

$$F = ([(\hat{\gamma}_1 - \hat{\gamma}_2) \sum_r \sum_c 1]/1)/((\epsilon_1^2 + \epsilon_2^2)/(2 \sum_r \sum_c 1 - 6)).$$

Under the hypothesis that $\gamma_1 = \gamma_2$, this statistic has an F distribution with 1,

$$2 \sum_r \sum_c 1 - 6$$

degrees of freedom. High values of F constitute a reason for rejecting the hypothesis that the constant surfaces are identical.

Rosenfeld *et al.* [4] suggested taking the products of $|\hat{\gamma}_1 - \hat{\gamma}_2|$ over varying size neighborhoods. Such a product will be large if each of the terms is high. High values for a set of neighborhood sizes will tend to arise in a region where the denominator of the F-statistic is low, thereby making the F statistic high for the largest neighborhood used. Of course, a better way of using the varying sized neighborhoods is to compute the F statistic itself for each of the various neighborhood sizes and then measure the strength of the step edge by the highest value F statistic computed.

There are few papers that approach edge detection statistically. Yakimovsky [5] suggests using a maximum likelihood principle for detecting step edges. His null hypothesis is that the sample values for the two adjacent neighborhoods come from the same normal distribution. The alternative hypothesis is that the distributions differ in mean and or variance. The disadvantage of Yakimovsky's test is that unless the sample is large, the distribution of the maximum likelihood test statistic is not known. Asymptotically however, -2 times the natural logarithm of the test statistic has a chi-squared distribution with 3 degrees of freedom. The F test discussed in Section 3 has the advantage of being an exact test under the assumption that the variances are identical. If the identical variance assumption is not appropriate, the maximum likelihood test could be extended to the case of the slope model.

Statistical approaches to edge detection require the use of a fitting model combined with a noise assumption to permit evaluating the significance of the edge strength measure. There seems to be few papers which do both. Besides Yakimovksy's maximum likelihood approach, Shanmugam *et al.* [20] derive an optimal edge filter whose behavior with noise is computed. Cohen and Toussaint [21] are concerned with the unequal distribution of random noise in the equal interval basis in a Hough transform space.

Nahi and Jahanshahi [15], Nahi and Assefi [12], Habibi [13], and Nahi [14] all use a statistical model of the object, background, and noise in a Bayesian and/or recursive estimation scheme to improve the image or estimate the boundary. The spatial models underlying these approaches are different from the fitting models discussed here. It would be interesting to find a method which combines both approaches.

As mentioned earlier, the fact that the slopes in two orthogonal directions determine the slope in any direction is well known in vector calculus. However, it seems not to be so well known in the image processing community. Prewitt [2], Kirsch [3] and Robinson [19] each suggested a set of masks which could be used to determine the gradient direction and magnitude for eight directions. Obviously, the masks of Prewitt, Kirsch, and Robinson not only do not give correct values for the gradient but also require two or four times more work to compute than the two masks required to determine $\hat{\alpha}$ and $\hat{\beta}$. Their application to determine the angle of step edges might seem more appropriate. However, O'Gorman [9] determined that for ideal straight-line step edges, the edge direction as determined by the direction of steepest slope using $\hat{\alpha}$ and $\hat{\beta}$ will not differ by more than 6.6° from the actual step edge direction. This implies that perhaps even for the problem of step edges, using many directional masks is not necessary.

6. BAYESIAN EDGE DETECTION AND REGION ANALYSIS

In Section 4 we have outlined a procedure for computing a statistic F whose distribution is known under the null hypothesis: no edge. This means that $P(F \mid \text{no edge})$ is known. An image of the computed F statistics can be histogrammed to estimate $P(F)$. Now if the prior probability of a nonedge were known, by Bayes' formula it is possible to determine the probability of no edge given the value of the statistic F:

$$P(\text{no edge} \mid F) = \frac{P(F \mid \text{no edge})\, P(\text{no edge})}{P(F)}.$$

Since there is either an edge or no edge, we have

$$P(\text{edge} \mid F) = 1 - P(\text{no edge} \mid F).$$

The image itself can tell us some about the prior probability of edge and nonedge. For example, under the assumption that all edges are of equal strength, it is clear that the histogram of the F-statistic image will be bimodal: one mode corresponding to the F statistics generated in no edge areas and the other mode corresponding to the F statistics generated in edge areas. The valley in the histogram gives us a threshold to detect edges and the probability on the side of the valley corresponding to the higher-valued F statistics gives us an estimate of the edge probability after observing the image.

We may have some prior knowledge about edge probability before observing the image and this probability can be used to help determine the edge probability after observing the image by biasing the choice of valley bottom. Also assuming a simple distributional form for $P(F \mid \text{no edge})$ and $P(F \mid \text{edge})$ it is possible to compute the priors $P(\text{no edge})$ and $P(\text{edge})$ by best fitting to the observed mixture $P(F)$ (see Haralick and Singh [18]).

There is no reason why the prior probabilities for edge and no edge cannot vary with position on the image. We can make these dynamic by making them a function of average neighborhood gray tone.

Consider, for example, a popular segmentation scheme which is appropriate when an image consists of a collection of objects which are constant in gray tone. In this case, the image histogram will have a tendency to be multimodal with the modes corresponding to the various gray tone shades of the objects. We may locate the valleys in the histogram or determine the parameters of the mixture distribution to find the valleys (Chow and Kaneko [17]) and use those points as decision boundaries in a pixel-by-pixel classification scheme. This scheme uses no spatial image structure and works poorly with moderate noise, producing a highly broken-up segmentation. But there is information in the location of the valleys. Small changes in gray tones in the valley regions should be more significant of an edge than similar-sized changes near one of the histogram modes. Knowing this, we may set the prior probability of an edge at each pixel to depend on the nearness of the average neighborhood gray tone to a valley in the histogram of the image grey tones.

From this perspective it is clear that the modal segmentation scheme corresponds to segmentation based on edge prior probabilities which are a function of gray tone. The simple F test of Section 3 corresponds to equal edge and nonedge prior probabilities.

7. CONCLUSION

We have discussed a surface fitting model by which edge detection and region delineation can be done. Our presentation has been theoretical. Future work will be empirical. We will be determining the validity of the model itself, the effectiveness of the statistical procedure for edge detection and region delineation, and possible ways of using the residual fitting error as a feature in its own right.

REFERENCES

1. L. G. Roberts, Machine perception of three-dimensional solids, in *Optical and Electroptical Information Processing* (J. T. Tippett, *et al.*, Eds.), pp. 159–197, MIT Press, Cambridge, Mass., 1965.
2. J. M. S. Prewitt, Object enhancement and extraction, in *Picture Processing and Psychopicotorics* (B. S. Lipkin and A. Rosenfeld, Eds.), pp. 75–149, Academic Press, New York, 1970.
3. R. Kirsch, Computer determination of the constituent structure of biological images, *Comput. Biomed. Res.*, **4**, 1971, 315–328.
4. A. Rosenfeld, Y. Lee, and R. Thomas, Edge and curve detection for texture discrimination, in *Picture Processing and Psychopictorics* (B. S. Lipkin and A. Rosenfeld, Eds.), pp. 381–393, Academic Press, New York, 1970.
5. Y. Yakimovsky, Boundary and object detection in real world images, *J. Assoc. Comput. Mach.* **23**, 1976, 599–618.
6. A. Rosenfeld and A. Kak, *Digital Picture Processing*, Academic Press, New York, 1976.
7. M. Brooks, Rationalizing edge detectors, *Computer Graphics Image Processing* **8**, 1978, 277–285.
8. F. O'Gorman and M. B. Clowes, Finding picture edges through collinearity of feature points, *IEEE Trans. Computers* **C-25**, April 1976, 449–456.
9. F. O'Gorman, Edge detection using Walsh functions, in *AISB Summer Conference*, 1976.
10. L. Meró and T. Vamos, Real-time edge-detection using local operators, in *3rd International Joint Conference on Pattern Recognition*, Coronado, California, November 1976.
11. M. Hueckel, An operator which locates edges in digital pictures, *J. Assoc. Comput. Mach.* **18**, 1971, 113–125.
12. N. E. Nahi and T. Assefi, Bayesian recursive image estimation, *IEEE Trans. Computers*, **C-21**, July 1972.
13. A. Habibi, Two-dimensional Bayesian estimate of images, *Proc. IEEE* **60**, July 1972.
14. N. E. Nahi, Role of recursive estimation in statistical image enhancement, *Proc. IEEE* **60**, July 1972.
15. N. E. Nahi and M. H. Jahanshahi, Image boundary estimation, *IEEE Trans. Computers* **C-26**, August 1977, 772–781.
16. R. Hummel, Feature detection using basic functions, *Computer Graphics Image Processing* **9**, 1979, 40–55.
17. C. K. Chow and T. Kaneko, Automatic boundary detection of the left ventricle from cineangiograms, *Comput. Biomed. Res.* **5**, 1972, 388–410.
18. R. Haralick and A. Singh, Boundary detection in images using a nonstationary statistical model, in *IEEE International Conference on Cybernetics and Society*, Washington, D.C., September 1977.
19. G. Robinson, Edge detection by compass gradient masks, *Computer Graphics Image Processing* **6**, 1977, 492–501.
20. K. S. Shanmugam, F. Dickey, and J. Green, An optimal frequency domain filter for edge detection in digital images, *IEEE Trans. Pattern Anal. Machine Intelligence* **PAMI-1**, January 1979, 39–47.
21. M. Cohen and G. Toussaint, On the detection of structures in noisy pictures, *Pattern Recognition* **9**, 1977, 95–98.